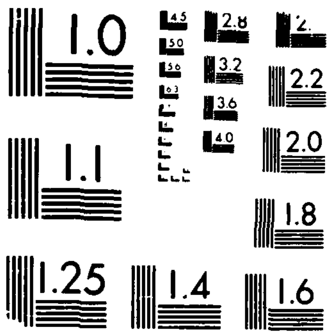
AD-A189 944 ALMOST SURE L(GAMMA)-NORM CONVERGENCE FOR DATA-BASED 1/1
HISTOGRAM DENSITY ES (U) PITTSBURGH UNIV PA CENTER FOR
MULTIVARIATE ANALYSIS L C ZHAO ET AL AUG 87 TR-87-30
UNCLASSIFIED AFOSR-TR-87-1843 F49620-85-C-0008 F/G 12/3 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

(2)

AFOSR-TR- 87-1843

Center for Multivariate Analysis

University of Pittsburgh

ALMOST SURE L_r-NORM CONVERGENCE FOR DATA-BASED HISTOGRAM DENSITY ESTIMATES

L. C. Zhao, P. R. Krishnaiah, X. R. Chen

Center for Multivariate Analysis
University of Pittsburgh
Pittsburgh PA 15260

August 1987

Technical Report Number 87-30

Center for Multivariate Analysis
University of Pittsburgh
Pittsburgh PA 15260

Research sponsored by the Air Force Office of Scientific Research (AFSC), under Contract F49620-85-C-0008. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

ALMOST SURE L_r-NORM CONVERGENCE FOR
DATA-BASED HISTOGRAM DENSITY ESTIMATES

L. C. Zhao, P. R. Krishnaiah, X. R. Chen

Center for Multivariate Analysis
University of Pittsburgh
Pittsburgh PA 15260

ABSTRACT

Let $X_1,\ldots,X_n$ be i.i.d. samples drawn from a d-dimensional distribution with density f. Partition the space R^d into a union of disjoint intervals $\{I_\ell = I(\ell,X_1,\ldots,X_n)\}$ with the form $I_\ell = \left\{x = (x^{(1)},\ldots,x^{(d)}): -\infty < a_{\ell i} \leq x^{(i)} \leq b_{\ell i} < \infty,\ i = 1,\ldots,d\right\}$. Define the data-based histogram estimate of f(x) based on this partition as

$$f_n(x) = \text{The number of } X_1,\ldots,X_n \text{ falling into } I_\ell$$
$$\div\ n \text{ times the volume of } I_\ell, \quad \text{for} \quad x \in I_\ell, \quad \ell = 1,2,\ldots$$

For given constant $r \geq 1$ we obtain the sufficient condition for $\lim_{n\to\infty} \int_{R^d} |f_n(x)-f(x)|^r dx = 0$. The results give substantial improvements upon the existing results.

AMS 1980 subject classifications: Primary 62F10; secondary 62H99.

Key words and phrases: Data-based, density estimator, empirical distribution, histogram.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 87-1843	2. GOVT ACCESSION NO. ADA189944	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Almost Sure L_r-Norm Convergence for Data-Based Histogram Density Estimates		5. TYPE OF REPORT & PERIOD COVERED ~~Technical~~ Journal Report - Sept 1987
		6. PERFORMING ORG. REPORT NUMBER 87-30
7. AUTHOR(s) L. C. Zhao, P. R. Krishnaiah, and X. R. Chen		8. CONTRACT OR GRANT NUMBER(s) F49620-85-C-0008
9. PERFORMING ORGANIZATION NAME AND ADDRESS Center for Multivariate Analysis, 515 Thackeray Hall, University of Pittsburgh, Pittsburgh, PA 15260		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304 A5
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research, Department of the Air Force, Bolling Air Force Base, DC 20332 NM		12. REPORT DATE ~~September~~ Aug 1987
		13. NUMBER OF PAGES 26
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Same as 11		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Key words and phrases: Data-based, density estimator, empirical distribution, histogram.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

DD FORM 1 JAN 73 1473

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Let $X_1,\dots,X_n$ be i.i.d. samples drawn from a d-dimensional distribution with density f. Partition the space R^d into a union of disjoint intervals $\{I_\ell = I(\ell,X_1,\dots,X_n)\}$ with the form $I_\ell = \{x = (x^{(1)},\dots,x^{(d)}): -\infty < a_{\ell i} \le x^{(i)} \le b_{\ell i} < \infty,\ i = 1,\dots,d\}$. Define the data-based histogram estimate of f(x) based on this partition as

$$f_n(x) = \text{The number of } X_1,\dots,X_n \text{ falling into } I_\ell \div n \text{ times the volume of } I_\ell, \quad \text{for } x \in I_\ell, \quad \ell = 1,2,\dots$$

For given constant $r \geq 1$ we obtain the sufficient condition for $\lim_{n\to\infty}\int_{R^d}|f_n(x)-f(x)|^r dx = 0$. The results give substantial improvements upon the existing results.

1. INTRODUCTION AND SUMMARY

Suppose that $X_1,\dots,X_n$ are i.i.d. samples of a d-dimensional random vector X. Throughout this paper, we shall denote by F the distribution of X, f the probability density function of X, $X^n = (X_1,\dots,X_n)$, and F_n the empirical distribution of X^n.

Let $f_n \equiv f_n(x) \equiv f_n(x;X^n)$ be an estimate of f based on X^n. For any constant $r \geq 1$, define

$$m_{nr} \equiv m_{nr}(X^n) \equiv \int |f_n(x)-f(x)|^r dx. \tag{1}$$

Here and in the sequel, $\int$ means $\int_{R^d}$. $m_{nr}^{1/r}$, to be called the L_r-norm of $f_n - f$, is a much - studied criterion in evaluating the performance of a density estimator. Quite a number of works have been done on the problem of convergence (to zero) of m_{nr} as the sample size n tends to infinity. We say that f_n is a L_r-norm consistent estimator of f if $m_{nr} \to 0$ as $n \to \infty$ in some sense.

For the kernel estimator

$$f_n(x) \equiv (nh_n^d)^{-1} \sum_{i=1}^{n} K(h_n^{-1}(x-X_i)),$$

where the kernel is assumed to be a probability density, Devroye [8] proved that the necessary and sufficient conditions for

$$\lim_{n\to\infty} m_{n1} = 0, \quad \text{a.s.} \tag{2}$$

are that $h_n \to 0$, $nh_n^d \to \infty$. Bai and Chen [3] solved the general case of $r \geq 1$, proving that the necessary and sufficient conditions for

$$\lim_{n\to\infty} m_{nr} = 0, \quad \text{a.s.} \quad \text{for some} \quad r \geq 1 \tag{3}$$

are that

$$h_n \to 0, \quad nh_n^d \to \infty, \quad \int f^r(x)dx < \infty. \quad \int K^r(u)du < \infty.$$

In the case of k_n-nearest neighbor estimator proposed by Loftsgarden and Quesenberry [10], Zhao Yue [12] proved that a sufficient condition for (3) in case of $r > 1$ is that

$$k_n/n \to 0, \quad k_n/\log n \to \infty, \quad \int f^r(x)dx < \infty \tag{4}$$

while the first and last in (4), and also that $k_n \to \infty$, are necessary for the truth of (3) $(r > 1)$.

Another important type of density-estimator is the histogram - ordinary histogram and data-based histogram. In ordinary histogram, the partitioning of range space of X, R^d, is done prior to the drawing of samples X^n. For this case, Abou-Jaoude [1], [2] (see also Devroye and Györfi [9], pp.19-23) obtained the necessary and sufficient conditions (imposed on the partition) for the truth of (2). Chen and Zhao [7] solved the general case of $r \geq 1$, for the particular partition

$$R^d = \bigcup_{k_1,\ldots,k_d=-\infty}^{\infty} \{x = (x^{(1)},\ldots,x^{(d)}): a_i + k_i h_n \leq x^{(i)} < a_i + (k_i+1)h_n,$$
$$1 \leq i \leq d\}.$$

Data-based histogram differs from the ordinary one in that the partition of space R^d defining the density estimate depends on the observations X^n. Thus, after obtaining X^n, we make a partition of R^d:

$$\Phi_n \equiv \{I(\ell, X^n): \ell = 1,2,\ldots\}$$

$$(\bigcup_{\ell=1}^{\infty} I(\ell, X^n) = R^d, \quad I(j,X^n) \cap I(k,X^n) = \emptyset \quad \text{when } j \neq k).$$

In this paper we consider only the case that $I(\ell,X^n)$, $\ell = 1,2,\ldots,$ are intervals in R^d of the form

$$[a_1,b_1)X[a_2,b_2)X\ldots X[a_d,b_d), \quad -\infty \le a_i < b_i \le \infty, \quad 1 \le i \le d.$$

For each $x \in R^d$, denote by $I_n(x)$ the unique interval in Φ_n containing x, and by $\lambda(I_n(x))$ the Lebesque measure of $I_n(x)$. The data-based histogram estimate f_n, based on the partition Φ_n, is defined by

$$f_n(x) \equiv F_n(I_n(x))/\lambda(I_n(x)). \tag{5}$$

For this estimate, the problem of L_r-norm consistency is much more complicated as compared with the ordinary histogram case. To begin with, for each positive integer n and positive constant t, denote by C_{nt} the number of intervals in $\Phi_n = \{I(\ell,X^n)\}$ fulfiling the condition

$$I(\ell,X^n) \cap \{x = (x^{(1)},\ldots,x^{(d)}) : |x^{(i)}| \le t, \quad 1 \le i \le d\} \neq \emptyset$$

and denote by $D(A)$ the diameter for any set $A \subseteq R^d$. Chen and Rubin [4] proved that

$$\lim_{n\to\infty} m_{n1} = 0, \quad \text{in probability} \tag{6}$$

under three conditions, two of them are:

$$\lim_{n\to\infty} D(I_n(x)) = 0, \quad \text{in probability, for } x \in R^d, \text{ a.e.}\lambda \tag{7}$$

$$\begin{cases} \lim_n C_{nt}/n = 0, & d = 1; \\ \lim_{n\to\infty} C_{nt}/\sqrt{n} = 0, & d > 1. \end{cases} \quad \text{in probability, for any } t > 0 \tag{8}$$

while the third one is of a rather complicated nature. Chen and Zhao [6] studied the strong consistency for the case of general d, proving the truth of (2) under the conditions:

$$\lim_{n\to\infty} D(I_n(x)) = 0, \quad \text{a.s.,} \quad \text{for} \quad x \in R^d, \quad \text{a.e.}\lambda \tag{7*}$$

$$\lim_{n\to\infty} C_{nt}\log n/n = 0, \quad \text{a.s.,} \quad \text{for any } t > 0 \tag{8*}$$

while Chen and Wang [5] obtained analogous result for this problem.

By comparing (8) and (8*) we see that although in case $d > 1$ (8*) is an improvement of (8), but in case $d = 1$, in achieving strong

Comparing the above two results, we see that in case $d > 1$ we have not only made the improvement by establishing a.s. convergence instead of convergence in probability, but also succeeded in some sense in weakening the conditions required, since (8*) requires a lower rate of convergence to zero than (8) - Of course, strictly speaking, (8) and (8*) are mutually exclusive. In case $d = 1$, in achieving strong consistency, we pay a price by requiring that c_{nt} is of the order $O(n/\log n)$ instead of $O(n)$. Motivated by the works of Devroye [8], Bai and Chen [3] and Chen and Zhao [7], we expect that the order $O(n)$ should be sufficient. In section 3, we shall prove that this is indeed the case:

THEOREM 1. Suppose that f_n is defined by (5), then (2) is true if (7*) and the following condition (9) are both true:

$$\lim_{n\to\infty} C_{nt}/n = 0, \quad \text{a.s.} \quad \text{for any} \quad t > 0 \tag{9}$$

The general case of $r > 1$ is considered in section 4. To formulate our result, for any interval $I = [a_1,b_1)X\ldots X[a_d,b_d)$ belonging to the partition Φ_n, write $a(I)$ for $\min_{1\le i\le d}(b_i-a_i)$. Also, for any $t > 0$, write

$Q_t = \{x = (x^{(1)},\ldots,x^{(d)}):|x^{(i)}| \le t,\ i = 1,\ldots,d\}$.

THEOREM 2. Suppose that f_n is defined by (5), then (3) is true if the following three conditions are satisfied:

$$\int f^r(x)dx < \infty \tag{10}$$

$$\sup\{D(I): I \in \Phi_n,\ I \cap Q_t \neq \emptyset\} \to 0 \quad \text{a.s.} \quad \text{for any } t > 0, \tag{11}$$

$$n(\inf_{I \in \Phi_n} a(I))^d \to \infty, \quad \text{a.s.} \tag{12}$$

The basic tool in our argument is an inequality establishing the exponential bound for the deviation between theoretical and empirical distributions over a class of partitions of R^d. The inequality is of independent interest and is the subject of section 2.

2. AN INEQUALITY

Suppose that μ is a probability measure on $\mathcal{B}^d$ - the σ- field of all Borel sets in R^d. Let $X_1,\ldots X_n,\ldots$ be i.i.d. random vectors with a common probability distribution μ, and μ_n be the empirical distribution of $X_1,\ldots,X_n$.

We call $\Phi \equiv \{A_1,\ldots,A_k\}$ a partition of R^d, if $A_1,\ldots, A_k$ are disjoint intervals of the form

$$[a_1,b_1)X\ldots X[a_d,b_d) \; : \; -\infty \le a_i < b_i \le \infty, \quad i = 1,\ldots,d$$

and $\bigcup_{i=1}^{k} A_i = R^d$. For fixed positive integer k, denote by $F \equiv F_k$ the collection of all such partitions, and define

$$D_n \equiv D_n(X^n) \equiv \sup\{\textstyle\sum_{A \in \Phi}|\mu_n(A)-\mu(A)| \; : \; \Phi \in F\}$$

It can easily be seen that there exists a countable subset $\{\Phi_i,\ i = 1, 2,\ldots\}$ $\subset F$, such that $D_n = \sup\{\Sigma_{A\in\Phi_i}|\mu_n(A) - \mu(A)|: i = 1,2,\ldots\}$, and $\{\Phi_i\}$ is independent of X^n. This shows that D_n is a random variable. We are now going to establish the following exponential bound for D_n:

THEOREM A. Given $\varepsilon \in (0,1)$, we have

$$P(D_n > \varepsilon) < 6 \exp(-n\varepsilon^2 2^{-9})$$

when $n \ge \max(k, 100\log 6/\varepsilon^2)$, and $(\frac{k}{n})\log(\frac{3en}{k}) < \varepsilon^2 2^{-9}(d+1)^{-1}$.

In proving the theorem, we borrow some idea from a work of Vapnik and Chervonenkis [11]. We also note that Theorem A extends a work of Devroye [8], which we formulate below as a lemma:

LEMMA 1 (Devroye [8]) Suppose that $R^d = \bigcup_{i=1}^{k} A_i$, $A_i \in \mathcal{B}^d$,

$i = 1,\ldots,k$, and $A_i \cap A_j = \emptyset$ when $i \neq j$. Then for given $\varepsilon > 0$ we have

$$P\left(\sum_{i=1}^{k} |\mu_n(A_i)-\mu(A_i)| \geq \varepsilon\right) \leq 3\exp(-n\varepsilon^2/25), \text{ when } k/n \leq \varepsilon^2/20$$

The following simple fact is also needed in the proof:

LEMMA 2. Let q_i, λ_i, $i = 1,\ldots,k$, be positive numbers. Write $a = \sum_{i=1}^{k} \lambda_i$, $b = \sum_{i=1}^{k}\lambda_i q_i$. We have

$$\prod_{i=1}^{k} q_i^{\lambda_i q_i} \geq \left(\frac{b}{a}\right)^b$$

and the equality holds if and only if $q_1 = \ldots = q_k$.

The proof is easy and therefore omitted.

Proof of the Theorem.

Write $X^{(n)} = (X_{n+1},\ldots,X_{2n})$, $X^{2n} = (X_1,\ldots,X_{2n})$, $\mu_n^* =$ the empirical measure of $X^{(n)}$, and

$$D_n(X^n,\Phi) = \sum_{A\in\Phi}|\mu_n(A)-\mu(A)|$$

$$D_n^*(X^{(n)},\Phi) = \sum_{A\in\Phi}|\mu_n^*(A)-\mu(A)|$$

$$G_n(X^{2n},\Phi) = \sum_{A\in\Phi}|\mu_n(A)-\mu_n^*(A)|$$

$$G_n \equiv G_n(X^{2n}) = \sup\{G_n(X^{2n},\Phi) : \Phi \in F\}$$

Since $\{G_n > \frac{\varepsilon}{2}\} = \bigcup_{i=1}^{\infty} \{G_n(\Phi_i) > \frac{\varepsilon}{2}\} \supset \bigcup_{i=1}^{\infty}\{D_n(\Phi_i) > \varepsilon\} \cap \{D_n^*(\Phi_i) < \varepsilon/2\}$ and $\{D_n(\Phi_i): i = 1,2,\ldots\}$, $\{D_n^*(\Phi_i): i = 1,2,\ldots\}$ are independent, it is well known that

$$P(G_n > \tfrac{\varepsilon}{2}) \geq \inf_i P(D_n^*(\Phi_i) < \varepsilon/2)P\left(\bigcup_{i=1}^{\infty}\{D_n(\Phi_i) > \varepsilon\}\right)$$
$$= \inf_i P(D_n^*(\Phi_i) < \varepsilon/2)P(D_n > \varepsilon)$$

Suppose that n satisfies the conditions indicated in Theorem A, then $k/n < \varepsilon^2/80$, and by Lemma 1 we have, simultaneously for all X^n:

$$P(D_n^*(X^{(n)},\Phi_i) \geq \varepsilon/2|X^n) \leq 3\exp(-n\varepsilon^2/100) \leq 1/2, \quad i = 1,2,\ldots$$

Therefore, $P(D_n^*(\Phi_i) < \varepsilon/2, \quad i = 1,2,\ldots,$ and

$$P(G_n > \tfrac{\varepsilon}{2}) \geq \tfrac{1}{2}P(D_n > \varepsilon) \tag{13}$$

From (13), it is seen that the proof of Theorem A reduces to the problem of finding an upper bound for $P(G_n > \varepsilon/2)$. For this purpose, denote by T a permutation $(j_1,j_2,\ldots,j_{2n})$ of $(1,2,\ldots,2n)$, so that $TX^{2n} = (X_{j_1},X_{j_2},\ldots,X_{j_{2n}})$. Further, denote by $\mu_n^{(T)}$ and $\mu_n^{(T)*}$ the empirical measures generated by $(X_{j_1},\ldots,X_{j_n})$ and $(X_{j_{n+1}},\ldots,X_{j_{2n}})$, respectively. Then it is readily seen that

$$P(G_n > \tfrac{\varepsilon}{2}) = \int_{R^{2nd}} \frac{1}{(2n)!} \sum_T I\{\sup_{\Phi\in F}\textstyle\sum_{A\in\Phi}|\mu_n^{(T)}(A) - \mu_n^{(T)*}(A)| > \tfrac{\varepsilon}{2}\}dP$$

$$\leq \int_{R^{2nd}} \frac{2}{(2n)!} \sum_T I\{\sup_{\Phi\in F}\textstyle\sum_{A\in\Phi}|\mu_n^{(T)}(A)-\mu_{2n}(A)| > \tfrac{\varepsilon}{4}\}dP \tag{14}$$

where the summation $\sum_T$ is taken over all (2n)! permutations of $(1,2,\ldots,2n)$, and $P = \mu^\infty$.

Now fix X^{2n}, and denote by U the set with elements $X_1,\ldots,X_{2n}$. Each $\Phi \in F$ induces a partition of the set U. Denote by $m_n(U)$ the number of different partitions induced by all $\Phi \in F$. We have

$$m_n(U) \leq \binom{2n+k-1}{k-1}^d \leq \binom{3n}{k-1}^d \leq \left(\frac{(3n)^k}{k!}\right)^d \leq \left(\frac{3en}{k}\right)^{kd} \tag{15}$$

Let F^* be a subset of F having $m_n(U)$ members, such that if $\Phi_i \in F^*$, $i = 1,2$ and $\Phi_1 \neq \Phi_2$, then Φ_1 and Φ_2 induce different partitions of U. We have

$$\frac{1}{(2n)!} \sum_T I\{\sup_{\Phi\in F} \sum_{A\in\Phi} |\mu_n^{(T)}(A)-\mu_{2n}(A)| > \frac{\varepsilon}{4}\}$$

$$= \frac{1}{(2n)!} \sum_T \sup_{\Phi\in F} I\{\sum_{A\in\Phi} |\mu_n^{(T)}(A) - \mu_{2n}(A)| > \frac{\varepsilon}{4}\}$$

$$= \frac{1}{(2n)!} \sum_T \sup_{\Phi\in F^*} I\{\sum_{A\in\Phi} |\mu_n^{(T)}(A)-\mu_{2n}(A)| > \frac{\varepsilon}{4}\}$$

$$\leq \sum_{\Phi\in F^*} \frac{1}{(2n)!} \sum_T I\{\sum_{A\in\Phi} |\mu_n^{(T)}(A)-\mu_{2n}(A)| > \frac{\varepsilon}{4}\}$$

$$\leq m_n(U) \sup_{\Phi\in F} \frac{1}{(2n)!} \sum_T I\{\sum_{A\in\Phi} |\mu_n^{(T)}(A)-\mu_{2n}(A)| > \frac{\varepsilon}{4}\} \tag{16}$$

Fix $\Phi = \{A_1,\dots,A_k\} \in F$. Denote by $Y_1,\dots,Y_n$ a random sample taken from U without replacement, and $\{Z_i,\ i \geq 1\}$ be a sequence of random samples taken from U with replacement. Write $\tilde{P}(\cdot) = P(\cdot|X^{2n})$, $\tilde{E}(\cdot) = E(\cdot|X^{2n})$, and

$$p_\ell = \mu_{2n}(A_\ell), \quad N_\ell = 2n\,p_\ell, \quad \ell = 1,\dots,k$$

$$V_n = \sum_{\ell=1}^k |I(\sum_{i=1}^n I(Y_i\in A_\ell)-np_\ell|, \quad W_n = \sum_{\ell=1}^k |\sum_{i=1}^n I(Z_i\in A_\ell)-np_\ell| \tag{17}$$

Then we have

$$\frac{1}{(2n)!} \sum_T I\{\sum_{A\in\Phi} |\mu_n^{(T)}(A)-\mu_{2n}(A)| > \frac{\varepsilon}{4}\}$$

$$= \tilde{E}\{I(\sum_{\ell=1}^k |\frac{1}{n} \sum_{i=1}^n I(Y_i\in A_\ell) - p_\ell| > \frac{\varepsilon}{4})\} = \tilde{P}(V_n > \varepsilon n/4) \tag{18}$$

Now we proceed to show that

$$\tilde{E}\{\exp(tV_n)\} \le (4\pi e^{1/6} n/k)^{k/2} \tilde{E}\{\exp(tW_n)\} \tag{19}$$

for any $t > 0$. In fact,

$$\tilde{E}\exp(tW_n)$$

$$= \sum{}' \frac{n!}{n! \dots n_k!} p_1^{n_1} \dots p_k^{n_k} \exp\{t \sum_{\ell=1}^{k} |n_\ell - np_\ell|\}, \tag{20}$$

where the summation Σ' is taken over all integer-valued vectors $(n_1, \dots, n_k)$ satisfying

$$n_1 \ge 0, \dots, n_k \ge 0 \quad \text{and} \quad \sum_{\ell=1}^{k} n_\ell = n.$$

In the same way, we have

$$\tilde{E}\{\exp(tV_n) = \Sigma' \binom{N_1}{n_1} \dots \binom{N_k}{n_k} \binom{2n}{n}^{-1} \exp\{t \sum_{\ell=1}^{k} |n_\ell - np_\ell|\}$$

$$= \Sigma' C(n_1, \dots, n_k) \frac{n!}{n_1! \dots n_k!} p_1^{n_1} \dots p_k^{n_k} \exp\{t \sum_{\ell=1}^{k} |n_\ell - np_\ell|\}. \tag{21}$$

Here, as usual, we put $\binom{n}{m} = 0$ for $m > n$. Also,

$$C(n_1, \dots, n_k) = \frac{n!(2n)^n}{(2n)!} \prod_{j=1}^{k} \left(\frac{N_j!}{(N_j - n_j)!} N_j^{-n_j}\right)$$

$$= \frac{n!(2n)^n}{(2n)!} \prod_{(I)} (N_j! N_j^{-N_j}) \prod_{(II)} \left(\frac{N_j!}{(N_j - n_j)!} N_j^{-n_j}\right),$$

where $\prod_{(I)}$ is taken over all j's satisfying $N_j = n_j$ and $\prod_{(II)}$ is taken over all j's satisfying $0 < n_j < N_j$. Using Stirling's formula

$$\sqrt{2\pi n}\, n^n e^{-n} < n! < \sqrt{2\pi n}\, n^n e^{-n+1/(12n)}$$

and the fact that $\sum_{j=1}^{k} n_j = n$, we get

$$C(n_1,\ldots,n_k)$$

$$\leq 2^{-n-1/2} e^{n+1/12} (2\pi)^{k/2} \exp(k/12-\sum_{j=1}^{k} n_j) \prod_{(II)} \left(\frac{N_j-n_j}{N_j}\right)^{-(N_j-n_j)-1/2}$$

$$\leq 2^{-n-1/2} (2\pi)^{k/2} e^{(k+1)/12} \prod_{(III)} \left(\frac{N_j-n_j}{N_j}\right)^{-(N_j-n_j)} \prod_{j=1}^{k} \sqrt{N_j}. \tag{22}$$

$\prod_{(III)}$, and the summation $\sum_{(III)}$ appearing below, are taken over all j's satisfying $0 \leq n_j < N_j$. Putting $q_j = (N_j-n_j)/N_j$, $\lambda_j = N_j$ in Lemma 2, we get

$$a \triangleq \sum_{(III)} \lambda_j = \sum_{(III)} N_j \leq 2n,$$

$$b \triangleq \sum_{(III)} \lambda_j q_j = \sum_{(III)} (N_j-n_j) = \sum_{j=1}^{k} (N_j-n_j) = n,$$

and

$$\prod_{(III)} \left(\frac{N_j-n_j}{N_j}\right)^{N_j-n_j} = \prod_{(III)} q_j^{\lambda_j q_j} \geq (b/a)^b \geq 2^{-n}. \tag{23}$$

On the other hand,

$$\prod_{j=1}^{k} \sqrt{N_j} \leq \left(\frac{1}{k} \sum_{j=1}^{k} N_j\right)^{k/2} = (2n/k)^{k/2}. \tag{24}$$

By (22) - (24),

$$C(n_1,\ldots,n_k) \leq 2^{-1/2} e^{(k+1)/12} (4\pi n/k)^{k/2} \leq (4\pi n/k)^{k/2} e^{k/12}, \tag{25}$$

and (19) follows from (20), (21) and (25).

Let N be a Poisson random variable, $E(N) = n$, and $N, (Y_1,\ldots,Y_n, Z_1,Z_2,\ldots)$ are independent. Since $\sum_{i=1}^{N} I(Z_i \in A_\ell)$, $\ell = 1,\ldots,k$, are independent Poisson variables with mean $np_1,\ldots,np_k$ respectively, it follows from (19) that for any $t > 0$,

$$\tilde{P}\{V_n > n\varepsilon/4\}$$

$$\leq \tilde{P}\{(N-n) > n\varepsilon/8\} + e^{-tn\varepsilon/4}\tilde{E}\{e^{tV_n}I(|N-n| \leq n\varepsilon/8)\}$$

$$= \tilde{P}\{|N-n| > n\varepsilon/8\} + e^{-tn\varepsilon/4}\tilde{E}e^{tV_n}\tilde{P}\{|N-n| \leq n\varepsilon/8\}$$

$$\leq \tilde{P}\{|N-n| > n\varepsilon/8\} + (4\pi e^{1/6}n/k)^{k/2}e^{-tn\varepsilon/4}\tilde{E}e^{tW_n}\tilde{P}\{|N-n| \leq n\varepsilon/8\}$$

$$= \tilde{P}\{|N-n| > n\varepsilon/8\} + (4\pi e^{1/6}n/k)^{k/2}e^{-tn\varepsilon/4}\tilde{E}\{e^{tW_n}I(|N-n| \leq n\varepsilon/8\}.$$

From the independence mentioned above and

$$e^{tW_n}I(|N-n| \leq n\varepsilon/8) \leq \exp\{t\sum_{\ell=1}^{k}|\sum_{i=1}^{N} I(Z_i \in A_\ell) - nP_\ell| + tn\varepsilon/8\},$$

it follows that

$$\tilde{P}\{V_n > n\varepsilon/4\}$$

$$\leq \tilde{P}\{|N-n| > n\varepsilon/8\}$$

$$+ (4\pi e^{1/6}n/k)^{k/2}e^{-tn\varepsilon/8}\prod_{\ell=1}^{k}\tilde{E}\{\exp(t|\sum_{i=1}^{N} I(Z_i \in A_\ell) - nP_\ell|)\}. \tag{26}$$

Now suppose that V is a Poisson variable and $EV = \lambda$. From $e^{-t} + t \leq e^t - t$ for $t > 0$, it follows that

$$E(e^{t|V-\lambda|}) \leq E\{e^{t(V-\lambda)} + e^{-t(V-\lambda)}\}$$

$$= \exp\{\lambda(e^t-1-t)\} + \exp\{\lambda(e^{-t}-1+t)\} \leq 2\exp\{\lambda(e^t-1-t)\}.$$

So we have

$$P\{|V-\lambda| \geq \lambda\varepsilon\} \leq E\{\exp(t|V-\lambda|-t\lambda\varepsilon)\}$$

$$\leq 2\exp\{-t\lambda\varepsilon + \lambda(e^t-1-t)\}.$$

Take $t = \log(1+\varepsilon)$, we get

$$P\{|V-\lambda| \geq \lambda\varepsilon\} \leq 2\exp\{\lambda(\varepsilon-(1+\varepsilon)\log(1+\varepsilon))\}$$

$$\leq 2\exp\{-\lambda\varepsilon^2/(2+2\varepsilon)\} \leq 2\exp(-\lambda\varepsilon^2/4)$$

for $\varepsilon e(0,1)$. Repeat this argument and take $t = \log(1+\varepsilon/8)$, by (26) we have

$$\tilde{P}\{V_n > n\varepsilon/4\}$$

$$\leq 2\exp(-n\varepsilon^2/256) + (4\pi e^{1/6}n/k)^{k/2}e^{-tn\varepsilon/8}\prod_{\ell=1}^{k}\{2\exp(nP_\ell(e^t-1-t))\}$$

$$\leq 2\exp(-n\varepsilon^2/256) + (4\pi e^{1/6}n/k)^{k/2}\, 2^k\exp\{n(e^t-1-t-t\varepsilon/8)\}$$

$$\leq 2\exp(-n\varepsilon^2/256) + (16\pi e^{1/6}n/k)^{k/2}\exp(-n\varepsilon^2/256)$$

$$\leq 3(16\pi e^{1/6}n/k)^{k/2}\exp(-n\varepsilon^2/256). \tag{27}$$

From (14) - (18) and (27), it follows that

$$P\{G_n > \varepsilon/2\} \leq 3(3en/k)^{kd}(16\pi e^{1/6}n/k)^{k/2}\exp(-n\varepsilon^2/256)$$

$$= 3\exp\{-n\varepsilon^2/256 + kd\log(3en/k) + \frac{k}{2}\log(16\pi e^{1/6}n/k)\}.$$

Under the conditions of Theorem A, $n/k > 16\, e^{1/6}/(9e^2)$ and $k(d+1)\log(3en/k) < n\varepsilon^2/2^9$. Hence,

$$P\{G_n > \varepsilon/2\} \leq 3\exp\{-n\varepsilon^2/256 + k(d+1)\log(3en/k)\}$$

$$\leq 3\exp\{-n\varepsilon^2/2^9\}.$$

From this and (13), Theorem A follows.

3. PROOF OF THEOREM 1

Define

$$\tilde{f}_n(x) = \int_{I_n(x)} f(u)du/\lambda(I_n(x)). \tag{28}$$

It is enough to show that for any $t > 0$,

$$\lim_{n\to\infty} \int_{Q_t} |f(x) - \tilde{f}_n(x)|dx = 0 \quad \text{a.s.} \tag{29}$$

and

$$\lim_{n\to\infty} \int_{Q_t} |f_n(x) - \tilde{f}_n(x)|dx = 0 \quad \text{a.s.} \tag{30}$$

For any $\varepsilon > 0$, we can find a function $g(x) \geq 0$ which is continuous on R^d and has a bounded support, such that $\int |f-g|dx < \varepsilon$. Define

$$\tilde{g}_n(x) = \int_{I_n(x)} g(u)du/\lambda(I_n(x)).$$

Then

$$\int_{Q_t} |f-\tilde{f}_n|dx \leq \int |f-g|dx + \int |\tilde{f}_n - \tilde{g}_n|dx + \int_{Q_t} |\tilde{g}_n - g|dx$$

$$\leq 2\int |f-g|dx + \int_{Q_t} |\hat{g}_n - g|dx$$

$$< 2\varepsilon + \int_{Q_t} |\tilde{g}_n - g|dx. \tag{31}$$

By (7*), there exists a set $B_0 \subset (R^d)^\infty$ such that $P(B_0) = 0$ and for $\omega \triangleq (X_1, X_2, \ldots) \overline{\in} B_0$ we have $\lim_{n\to\infty} D(I_n(x)) = 0$ for $x \in R^d$, a.e.λ, and in turn it follows that $\lim_{n\to\infty} \tilde{g}_n(x) = g(x)$ for $x \in R^d$, a.e.λ.

By the dominated convergence theorem,

$$\lim_{n\to\infty} \int_{Q_t} |\tilde{g}_n(x) - g(x)|dx = 0 \quad \text{a.s.}, \tag{32}$$

and (29) follows from (31) and (32).

From (9) it can be shown that there exists a sequence $\{\delta_n\}$ of positive numbers such that $\lim_{n\to\infty} \delta_n = 0$ and

$$\lim_{n\to\infty} C_{nt}/[n\delta_n] = 0 \quad \text{a.s.},$$

where $[n\delta_n]$ denotes the integer part of $n\delta_n$. For any $\varepsilon \in (0,1)$, there exists a set $B_{1-\varepsilon/2} \subset (R^d)^\infty$ such that $P(B_{1-\varepsilon/2}) > 1 - \varepsilon/2$ and

$$\lim_{n\to\infty} C_{nt}/[n\delta_n] = 0 \quad \text{uniformly for } (X_1, X_2, \ldots) \in B_{1-\varepsilon/2}.$$

So there exists a positive integer N such that

$$C_{nt} < [n\delta_n], \quad \text{for } n \geq N \text{ and } (X_1, X_2, \ldots) \in B_{1-\varepsilon/2}.$$

Now we recall $\Phi_n \equiv \Phi_n(X^n) = \{I(\ell, X^n), \quad \ell = 1,2,\ldots\}$ is the partition of R^d which defines the data-based histogram f_n. It is easy to see that we can find $k \leq 3^d C_{nt}$ and $\Phi \in F_k$ such that

$$\{I : I \in \Phi, I \cap Q_t \neq \emptyset\} = \{I : I \in \Phi_n, I \cap Q_t \neq \emptyset\}.$$

Hence, for $(X_1, X_2, \ldots) \in B_{1-\varepsilon/2}$, $n \geq N$ and $k = 3^d[n\delta_n]$, we have

$$\begin{aligned} \int_{Q_t} |f_n(x) - \tilde{f}_n(x)|dx &\leq \Sigma_{I\in\Phi_n, I\cap Q_t \neq \emptyset} |F_n(I) - F(I)| \\ &\leq \sup_{\Phi\in F_k} \Sigma_{A\in\Phi} |F_n(A) - F(A)| \triangleq D_n. \end{aligned} \tag{33}$$

Since $k/n = 3^d[n\delta_n]/n \leq 3^d\delta_n \to 0$ as $n \to \infty$, from Theorem A, we have

$$\lim_{n\to\infty} D_n = 0 \quad \text{a.s.} \tag{34}$$

By (33) and (34), there exists a set $B_{1-\varepsilon} \subset (R^d)^\infty$ such that $B_{1-\varepsilon} \subset B_{1-\varepsilon/2}$, $P(B_{1-\varepsilon}) > 1 - \varepsilon$ and

$$\lim_{n\to\infty} \int_{Q_t} |f_n(x) - \tilde{f}_n(x)| dx = 0 \quad \text{uniformly for} \quad (X_1, X_2, \ldots) \in B_{1-\varepsilon}.$$

Since $\varepsilon > 0$ is arbitrarily given, (30) is proved, and the proof of Theorem 1 is completed.

4. PROOF OF THEOREM 2

Define $\tilde{f}_n(x)$ as before by (28). Find a nonnegative function g, continuous everywhere on R^d and with a bounded support, such that $\int|f-g|^r dx < \varepsilon^r$. Put

$$\tilde{g}_n(x) = \int_{I_n(x)} g(u)du/\lambda(I_n(x)).$$

Then

$$\begin{aligned}(\int|f-\tilde{f}_n|^r dx)^{1/r} &\le (\int|f-g|^r dx)^{1/r} + (\int|\tilde{f}_n-\tilde{g}_n|^r dx)^{1/r} + (\int|\tilde{g}_n-g|^r dx)^{1/r} \\ &\le 2(\int|f-g|^r dx)^{1/r} + (\int|\tilde{g}_n-g|^r dx)^{1/r} \\ &< 2\varepsilon + (\int|\tilde{g}_n-g|^r dx)^{1/r}. \end{aligned} \tag{35}$$

By (11), for any $x \in R^d$, we have

$$D(I_n(x)) \to 0 \quad \text{a.s.}$$

There exists a set $B_0 \subset (R^d)^\infty$ such that $P(B_0) = 0$ and for $\omega \equiv (X_1,X_2,\ldots) \bar{\in} B_0$ we have $\lim_{n\to\infty} D(I_n(x)) = 0$ for $x \in R^d$, a.e.λ, and in turn it follows that $\lim_{n\to\infty}\tilde{g}_n(x) = g(x)$ for $x \in R^d$, a.e.λ. By the dominated convergence theorem,

$$\lim_{n\to\infty} \int|\tilde{g}_n-g|^r dx = 0 \quad \text{for} \quad \omega \bar{\in} B_0. \tag{36}$$

By (35) and (36), we have

$$\lim_{n\to\infty} \int|f(x)-\tilde{f}_n(x)|dx = 0 \quad \text{a.s.} \tag{37}$$

Now we proceed to prove that

$$\int|f_n-\tilde{f}_n|^r dx = \sum_{I_\ell \in \Phi_n} |F_n(I_\ell) - F(I_\ell)|^r/\lambda(I_\ell)^{r-1} \to 0, \quad \text{a.s.} \tag{38}$$

Put $H = \inf_{I \in \Phi_n} a(I)$. Since $nH^d \to \infty$ a.s., it can be shown that there exists a sequence $C_n \to \infty$ such that $\lim_{n\to\infty} nH^d/C_n^d = \infty$ a.s.. Without loss of generality, we can assume that $C_n^d/n \to 0$. Take $h = h_n = C_n/n^{1/d}$, then $h_n \to 0$, $nh_n^d \to \infty$ and $H/h_n \to \infty$, a.s.

Construct a partition of R^d into disjoint finite intervals, say $\Psi_n = \{\Delta_1, \Delta_2, \ldots\}$, where Δ_m's are all cubes with the same edge length h. Define

$$\xi_n(x) = F_n(\Delta_m)/h^d \quad \text{for} \quad x \in \Delta_m, \quad m = 1,2,\ldots$$

and

$$\tilde{\xi}_n(x) = F(\Delta_m)/h^d \quad \text{for} \quad x \in \Delta_m, \quad m = 1,2,\ldots$$

By the theorem of [7],

$$\lim_{n\to\infty} \int |\xi_n(x)-f(x)|^r dx = 0. \quad \text{a.s.}$$

An argument similar to that leading to (37) gives

$$\lim_{n\to\infty} \int |\tilde{\xi}_n(x)-f(x)|^r dx = 0.$$

So we have

$$\int |\xi_n(x)-\tilde{\xi}_n(x)|^r dx = \sum_{\Delta_m \in \Psi_n} |F_n(\Delta_m)-F(\Delta_m)|^r/(h^d)^{r-1}$$

$$\to 0 \quad \text{a.s.} \quad \text{as} \quad n \to \infty. \tag{39}$$

For $I_\ell \in \Phi_n$, denote by $H_{1\ell}, \ldots, H_{d\ell}$ the lengths of the edges of I_ℓ, and write

$$M_\ell = \{m : \Delta_m \in \Psi_n, \ \Delta_m \subset I_\ell\},$$

$$\tilde{M}_\ell = \{m : \Delta_m \in \Psi_n, \ \Delta_m \cap I_\ell \neq \emptyset, \ \Delta_m \setminus I_\ell \neq \emptyset\}. \tag{40}$$

Since $H/h_n \to \infty$ a.s., we can find $B_* \subset (R^d)^\infty$ such that $P(B_*) = 0$ and $H/h_n \to \infty$ for $\omega \bar{\in} B_*$. In the sequel we always keep $\omega \bar{\in} B_*$. Thus, for n large, $H_{i\ell} > 2h$ for all i and ℓ. We have

$$\prod_{i=1}^{d}(H_{i\ell}/h - 2) \leq \#(M_\ell) \leq \prod_{i=1}^{d}(H_{i\ell}/h), \tag{41}$$

and

$$\begin{aligned}
\#(\tilde{M}_\ell) &\leq \prod_{i=1}^{d}(H_{i\ell}/h + 2) - \prod_{i=1}^{d}(H_{i\ell}/h - 2) \\
&\leq \lambda(I_\ell)h^{-d}\{\prod_{i=1}^{d}(1+2h/H_{i\ell}) - \prod_{i=1}^{d}(1-2h/H_{i\ell})\} \\
&\leq \lambda(I_\ell)h^{-d}\{(1+2h/H)^d - (1-2h/H)^d\} \\
&\leq h^{-d}\lambda(I_\ell)C(d)h/H,
\end{aligned} \tag{42}$$

where

$$C(d) = 2^d(2^{d-1}+1).$$

Now, by (41) and (39) we have

$$\begin{aligned}
\rho_n &\triangleq \sum_{I_\ell \in \Phi_n} \left|F_n\left(\sum_{m\in M_\ell}\Delta_m\right) - F\left(\sum_{m\in M_\ell}\Delta_m\right)\right|^r/\lambda(I_\ell)^{r-1} \\
&\leq \sum_{I_\ell \in \Phi_n}(\#(M_\ell))^{r-1}\sum_{m\in M_\ell}|F_n(\Delta_m) - F(\Delta_m)|^r/\lambda(I_\ell)^{r-1} \\
&\leq \sum_{\Delta_m \in \Psi_n}|F_n(\Delta_m) - F(\Delta_m)|^r/(h^d)^{r-1} \to 0, \quad \text{a.s.}
\end{aligned} \tag{43}$$

On the other hand, by (42) we have

$$\begin{aligned}
\tilde{\rho}_n &\triangleq \sum_{I_\ell \in \Phi_n}|F_n(\sum_{m\in\tilde{M}_\ell}(I_\ell\Delta_m)) - F(\sum_{m\in\tilde{M}_\ell}(I_\ell\Delta_m))|^r/\lambda(I_\ell)^{r-1} \\
&\leq \sum_{I_\ell \in \Phi_n}(\#(\tilde{M}_\ell))^{r-1}\sum_{m\in\tilde{M}_\ell}|F_n(I_\ell\Delta_m) - F(I_\ell\Delta_m)|^r/\lambda(I_\ell)^{r-1}
\end{aligned}$$

$$\le (hC(d)H^{-1})^{r-1} \sum_{I_\ell \in \Phi_n} \sum_{m \in \tilde{M}_\ell} |F_n(I_\ell \Delta_m) - F(I_\ell \Delta_m)|^r/(h^d)^{r-1}. \tag{44}$$

For each $\Delta_m \in \Psi_n$, define

$$N_m = \{\ell\colon\ I_\ell \in \Phi_n,\ I_\ell \cap \Delta_m \neq \emptyset\} \tag{45}$$

Since $H_{i\ell} > 2h$ for all i and ℓ, for any m the set N_m contains at most 2^d elements. By (44),

$$\begin{aligned}
\tilde{\rho}_n &\le (C(d)h/H)^{r-1} \sum_{\Delta_m \in \Psi_n} \sum_{\ell \in N_m} |F_n(I_\ell \Delta_m) - F(I_\ell \Delta_m)|^r/(h^d)^{r-1} \\
&\le (C(d)h/H)^{r-1} \sum_{\Delta_m \in \Psi_n} \#(N_m) 2^{r-1}(F_n(\Delta_m)^r + F(\Delta_m)^r/(h^d)^{r-1} \\
&\le 2^{d+r-1}(C(d)h/H)^{r-1} \sum_{\Delta_m \in \Psi_n} 2^{r-1}|F_n(\Delta_m) - F(\Delta_m)|^r/(h^d)^{r-1} \\
&\quad + 2^{d+r-1}(C(d)h/H)^{r-1} \sum_{\Delta_m \in \Psi_n} (2^{r-1}+1)F(\Delta_m)^r/(h^d)^{r-1} \\
&\triangleq \tilde{\rho}_{n1} + \tilde{\rho}_{n2}.
\end{aligned} \tag{46}$$

By (43),

$$\lim_{n\to\infty} \tilde{\rho}_{n1} = 0 \quad \text{a.s.} \tag{47}$$

By Jensen's inequality,

$$\begin{aligned}
\sum_{\Delta_m \in \Psi_n} F(\Delta_m)^r/(h^d)^{r-1} &= \sum_{\Delta_m \in \Psi_n} \left|h^{-d} \int_{\Delta_m} f(x)dx\right|^r h^d \\
&\le \sum_{\Delta_m \in \Psi_n} \int_{\Delta_m} f^r(x)dx = \int f^r(x)dx,
\end{aligned}$$

which implies that

$$\lim_{n\to\infty} \tilde{\rho}_{n2} = 0 \quad \text{a.s.} \tag{48}$$

From (46) - (48), we obtain

$$\lim_{n\to\infty} \tilde{\rho}_n = 0 \quad \text{a.s.} \tag{49}$$

By (43) and (49), we have

$$\int |f_n(x) - \tilde{f}_n(x)|^r dx = \sum_{I_\ell \in \Phi_n} |F_n(I_\ell) - F(I_\ell)|^r / \lambda(I_\ell)^{r-1}$$

$$\leq 2^{r-1}(\rho_n + \tilde{\rho}_n) \to 0 \quad \text{a.s.}$$

Thus, (38) is proved, and Theorem 2 follows from (37) and (38).

REFERENCES

[1] ABOU-JAOUDE, S. (1976). Sur une condition nécessaire et suffisante de L_1-convergence presque complète de l'estimator de la partition fixe pour une densité. Comptes Rendus de l'Académie des Science de Paris Série A 283. 1107 - 1110.

[2] ABOU-JAOUDE, S. (1976). Conditions nécessaires et suffisantes de convergence L_1 en probabilité de l'histogramme pour une densité. Annales de l'Institut Henri Poincaré 12. 213-231.

[3] BAI, Z. D., and CHEN, X. R., (1987). Necessary and sufficient conditions for the convergence of integrated and mean-integrated p-th order error of the kernel density estimators. Tech. Report 87-06. Center for Multivariate Analysis, University of Pittsburgh.

[4] CHEN, J., and RUBIN Herman, (1984). On the consistency property of the data-based histogram density estimator. Technical Report #84-11, Purdue University.

[5] CHEN, X. R., and WANG, S. R., (1985). L_1-norm strong consistency for data-based histogram estimates of multidimensional density. Personal communication.

[6] CHEN, X. R., and ZHAO, L. C., (1987). Almost sure L_1-Norm convergence for data-based histogram density estimates. Journal of Multivariate Analysis, 21,179-188.

[7] CHEN, X. R., and ZHAO, L. C., (1987). Necessary and sufficient conditions for the convergence of integrated and mean-integrated r-th order error of histogram estimates. Tech. Report No. 87-07, Center for Multivariate Analysis, University of Pittsburgh.

[8] DEVROYE, L., (1983). The equivalence of weak, strong and complete convergence in L_1 for kernel density estimates. Ann. Statist. 11, 896-904.

[9] DEVROYE, L., and GYÖRFI, L., (1985). Nonparametric Density Estimation: The L_1 View. John Wiley.

[10] LOFTSGARDEN, D. O., and QUESENBERRY, C. P., (1965). A nonparametric estimate of a multivariate probability density function. Ann. Math. Statist. 28, 1049-1051.

[11] VAPNIK, V. N., and CHERVONENKIS, A., (1971). On uniform convergence of the frequencies of events to their probabilities. Theor. Probability Appl., 16, 264-280.

[12] ZHAO, Yue (1986). Strong L_p-norm consistency of one-dimensional nearest neighbor density estimates.

END
DATE
FILMED
5-88
DTIC